

The 6th Answer Set Programming Competition

Martin Gebser, Marco Maratea, Francesco Ricca



13th International Conference on Logic Programming
and Non-monotonic Reasoning

Outline

- ① The Sixth ASP Competition
- ② Format and Setup
- ③ Participants and Results

The Sixth ASP Competition

An event back to the usual timeline

- One year after the FLoC Olympic Games
- Hosted by LPNMR
- Biennial event

The Sixth ASP Competition

An event back to the usual timeline

- One year after the FLoC Olympic Games
- Hosted by LPNMR
- Biennial event

Goals

- Measure the progress of the state of the art in ASP solving
- Improve benchmarks suite for robust evaluation
- Study the behavior of different solving techniques

The 6th Competition Setting

Improvements on the format

- Basic design choices maintained
- Some important novelties

The 6th Competition Setting

Improvements on the format

- Basic design choices maintained
- Some important novelties

Competition Setting

- System competition only and modeling competition on site
- Benchmark classification based on language features
- Benchmarks from past editions
 - The best encodings from 2014
 - Updated instance sets
 - New “real-world” benchmarks
- New instance selection process
- Updated versions of solvers, and newcomers

Outline

- 1 The Sixth ASP Competition
- 2 Format and Setup**
- 3 Participants and Results

System Competition Format

Sub tracks based on language features

Track 1 (Basic) normal LP + simple built-ins

Track 2 (Advanced) + choices, aggregates, HCF disjunction, query

Track 3 (Optimization) + weak constraints

Track 4 (Unrestricted) + non-HCF disjunction

Two Categories

- Single-Processor (restricted to 1-CPU Core)
- Multi-Processor (up to 8-CPU Cores)

System Competition Format

Sub tracks based on language features

Track 1 (Basic) normal LP + simple built-ins

Track 2 (Advanced) + choices, aggregates, HCF disjunction, query

Track 3 (Optimization) + weak constraints

Track 4 (Unrestricted) + non-HCF disjunction

Two Categories

- Single-Processor (restricted to 1-CPU Core)
- Multi-Processor (up to 8-CPU Cores)

Marathon ← **NEW!!**

- The best solver of each team
- Time limit extended by one order of magnitude
→ Assess solvers on hard instances

Setup

System Inputs

- Fixed input in ASP-Core-2
- Scripts run with fixed parameters
- Fixed encoding + instance from STD input

System Environment

- Debian Linux 64bit with Intel Xeon E5-4610v2 CPUs
- Time limits
 - Competition: 20 minutes
 - Marathon: 3 hours
- Memory Limit: 12 GB
- Multi-processor track: up to 8 cores (16 virtual CPUs)

Scoring

ASP Competition 2014 Scoring

- Consider number of solved instances for decision problems
- Rank solvers on optimization problems by solution quality
- Runtime for tiebreaking

Decision and Query Problems

$$\text{Score}(\text{Solver}, \text{Problem}) = \# \text{Solved}(\text{Solver}) * 5$$

Optimization Problems

$$\text{Score}(\text{Solver}, \text{Problem}) = \sum_{\text{Instances } I} \frac{\# \text{NotBetter}(\text{Solver}, I) * 5}{\# \text{Participants}}$$

Scoring

ASP Competition 2014 Scoring

- Consider number of solved instances for decision problems
- Rank solvers on optimization problems by solution quality
- Runtime for tiebreaking

Additional Criteria

- Problems are equally weighted up to 100 points each
- Incorrect answers: disqualification on per problem basis
- Final scores by summing over all problems

Benchmark Suite

Benchmarks from 2014

- Considered all the domains from 5th edition
- Selected the encoding variant that exhibited better performance in the 5th edition
- Updated instance sets for
 - Knight Tour with Holes, Stable Marriage,
 - Ricochet Robots, and Maximal Clique
- **Hardness-based classification of instances**
 - Inspired of SAT Competition
 - Exploiting best solvers from the 5th competition
 - Robust selection

Benchmark Suite: Domains from past editions

Domain	App	Problem	Encoding	
<i>Graph Colouring</i>		Decision	2014	Track #1
<i>Hanoi Tower</i>		Decision	2014	
<i>Knight Tour with Holes</i>		Decision	2014	
<i>Labyrinth</i>		Decision	2013	
<i>Stable Marriage</i>		Decision	2014	
<i>Visit-all</i>		Decision	2014	
<i>Bottle Filling</i>		Decision	2013	Track #2
<i>Graceful Graphs</i>		Decision	2013	
<i>Incremental Scheduling</i>	✓	Decision	2014	
<i>Nomystery</i>		Decision	2014	
<i>Partner Units</i>	✓	Decision	2014	
<i>Permutation Pattern Matching</i>		Decision	2014	
<i>Qualitative Spatial Reasoning</i>		Decision	2014	
<i>Reachability</i>		Query	2013	
<i>Ricochet Robots</i>		Decision	2013	
<i>Sokoban</i>		Decision	2014	
<i>Solitaire</i>		Decision	2014	
<i>Weighted-Sequence Problem</i>		Decision	2014	
<i>Connected Still Life*</i>		Optimization	2013	Track #3
<i>Crossing Minimization</i>	✓	Optimization	2014	
<i>Maximal Clique</i>		Optimization	2014	
<i>Valves Location</i>	✓	Optimization	2013	
<i>Abstract Dialectical Frameworks</i>		Optimization	2013	Track #4
<i>Complex Optimization</i>	✓	Decision	2014	
<i>Minimal Diagnosis</i>	✓	Decision	2014	
<i>Strategic Companies</i>		Query	2013	

Benchmark Suite: New domains

Domain	App	Problem	
<i>Combined Configuration</i>	✓	Decision	Tr. #2
<i>Consistent Query Answering</i>	✓	Query	
<i>MaxSAT</i>	✓	Optimization	Track #3
<i>Steiner Tree</i>	✓	Optimization	
<i>System Synthesis</i>	✓	Optimization	
<i>Video Streaming</i>	✓	Optimization	

Benchmark Classification (1)

Run the three best solvers of 5th ASP Comp

- clasp, lp2normal+clasp, wasp1.5
- same setting as competition
- 40 min TO (twice the timeout)

Some numbers

- 32 domains
- 5058 instances
- about 212 days of execution

Benchmark Classification (2)

- (non-groundable)** Instances that could not be grounded by any top-performing system within the timeout.
- (very easy)** Instances solved by all top-performing systems in less than 20 seconds.
- (easy)** Instances solved by all top-performing systems in less than 2 minutes.
- (medium)** Instances solved by all top-performing systems within the timeout.
- (hard)** Instances solved by at least one among the top-performing systems within 40 minutes.
- (too hard)** Instances that could not be solved (no solution produced in case of Optimization problems) by any of the top-performing systems within 40 minutes.

Instance Selection

Instance Selection (Criteria)

- 20 instances are included in each domain
- Exclude non-groundable instances
- Each class shall contribute 20% to each domain
- Discard domains mostly made of easy instances
- Balance satisfiable and unsatisfiable instances for decision
- Prefer satisfiable instances for optimization and query
- Random selection from each class + 20% totally random

Instance Selection

Instance Selection (Criteria)

- 20 instances are included in each domain
 - Exclude non-groundable instances
 - Each class shall contribute 20% to each domain
 - Discard domains mostly made of easy instances
 - Balance satisfiable and unsatisfiable instances for decision
 - Prefer satisfiable instances for optimization and query
 - Random selection from each class + 20% totally random
-
- **Selection implemented in ASP!!!!**
 - **Random seed:** the concatenation of winning numbers in the EuroMillions lottery of 23rd June 2015

Selection Statistics

Benchmark Suite

- 28 selected problems
- 4 too easy/uneven problems discarded
 - BottleFillingProblem
 - HanoiTower
 - Solitaire
 - Weighted-SequenceProblem
- 88 non-groundable instances
 - 86 IncrementalScheduling
 - 2 Sokoban
- Statistics about old solvers
 - Can be used for measuring the improvement of the state of the art

Outline

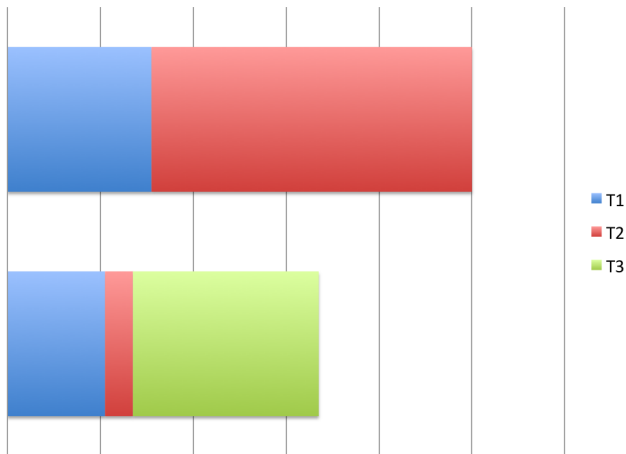
- 1 The Sixth ASP Competition
- 2 Format and Setup
- 3 Participants and Results

Participants

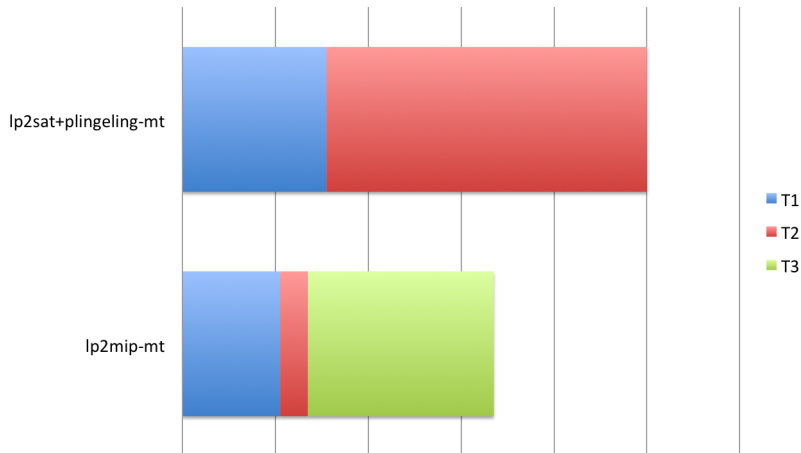
The competition featured 13 systems coming from three teams

- **Aalto Team**, Aalto University (9 solvers):
LP2SAT+LINGELING, LP2SAT+PLINGELING-MT, LP2ACYCASP+CLASP,
LP2ACYCPB+CLASP, LP2ACYCSAT+CLASP, LP2ACYCSAT+GLUCOSE,
LP2MIP, LP2MIP-MT, LP2NORMAL+CLASP
- **ME-ASP Team**, University of Genoa, University of Sassari,
University of Calabria (1 solver):
ME-ASP
- **Wasp Team**, University of Calabria (3 solvers):
WASP, WASP+DLV, JWASP

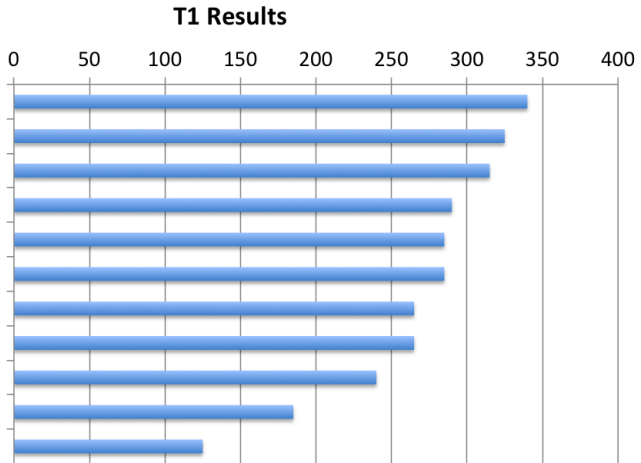
Results: Multi Processor Track



Results: Multi Processor Track

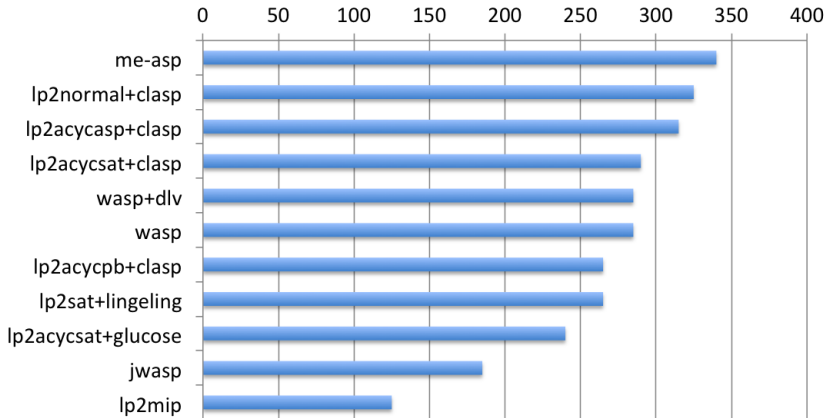


Results: Track 1 - Basic

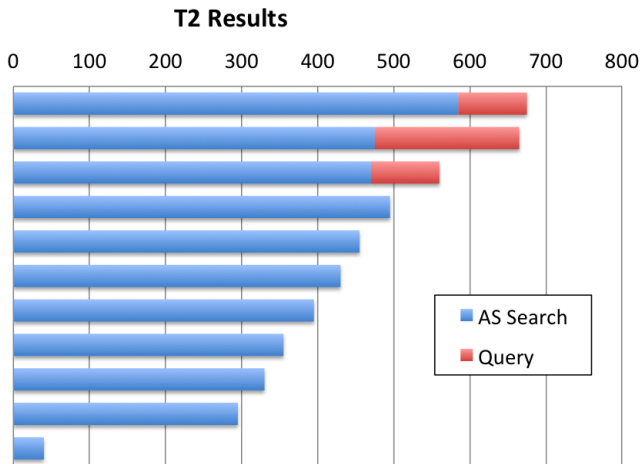


Results: Track 1 - Basic

T1 Results

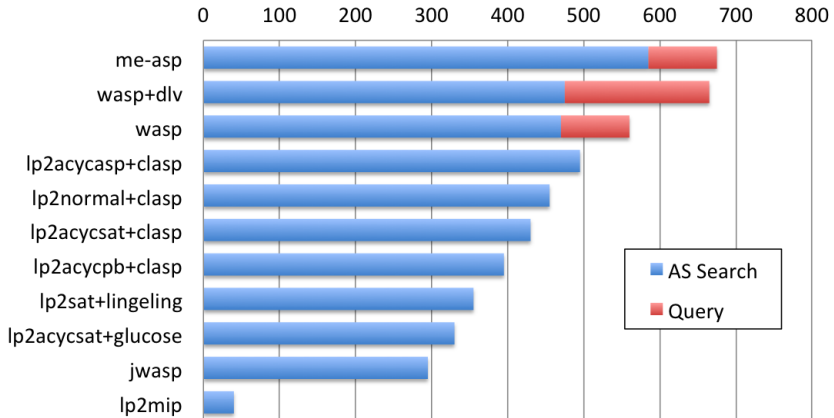


Results: Track 2 - Advanced

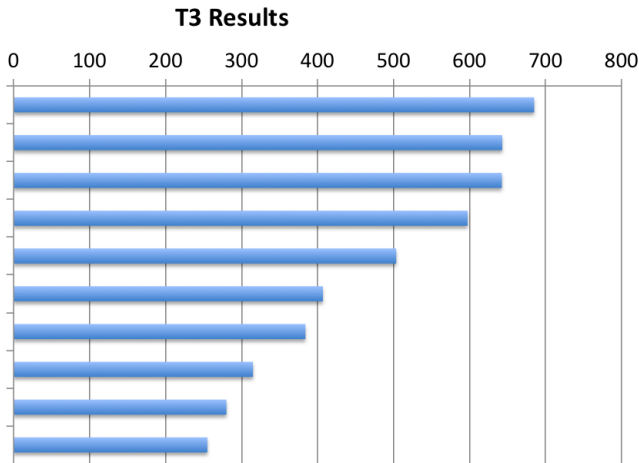


Results: Track 2 - Advanced

T2 Results

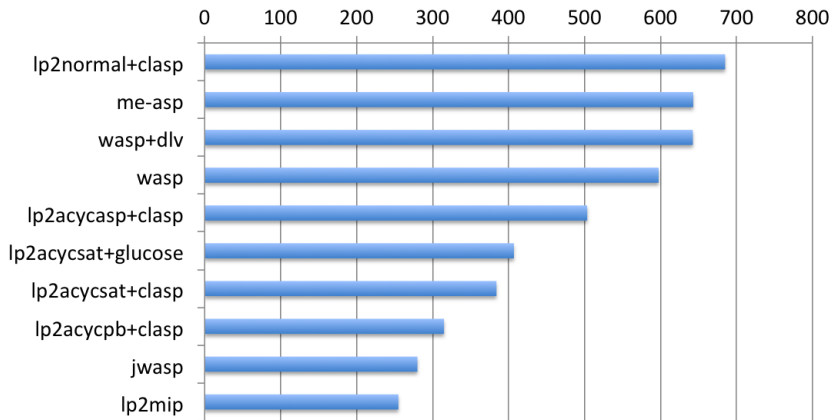


Results: Track 3 - Optimization

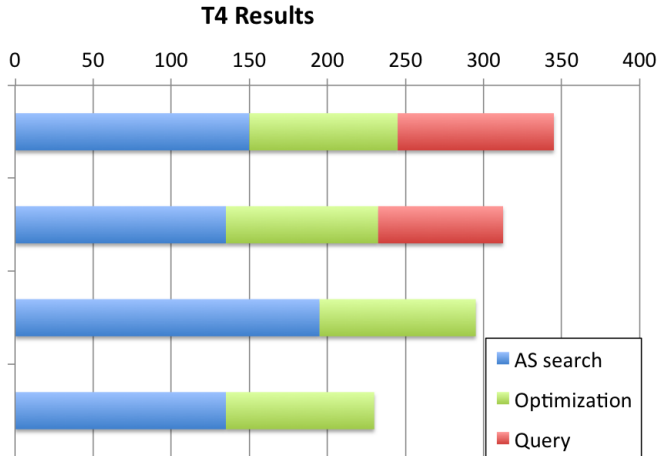


Results: Track 3 - Optimization

T3 Results

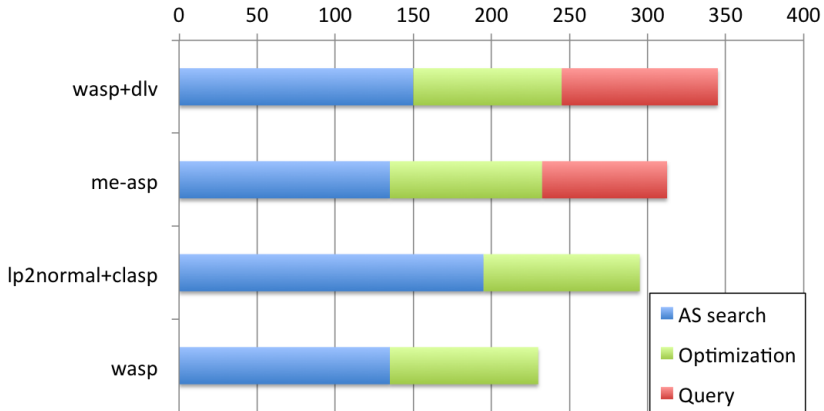


Results: Track 4 - Unrestricted

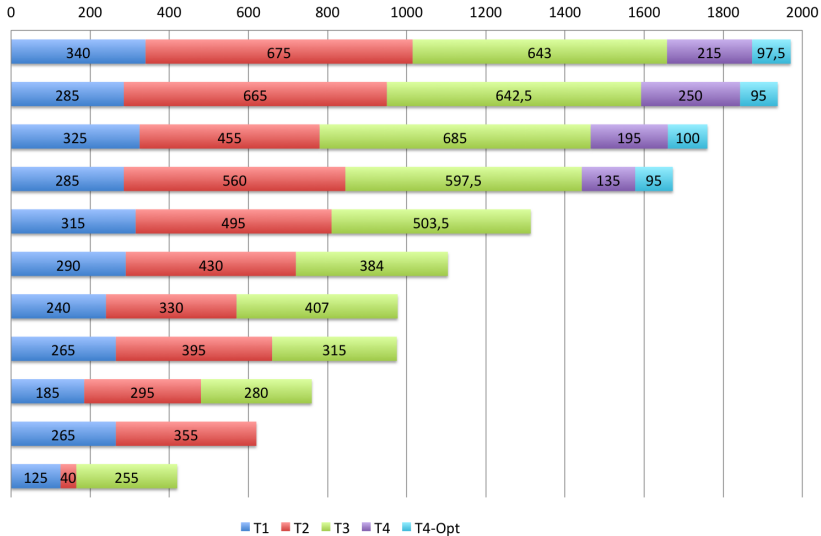


Results: Track 4 - Unrestricted

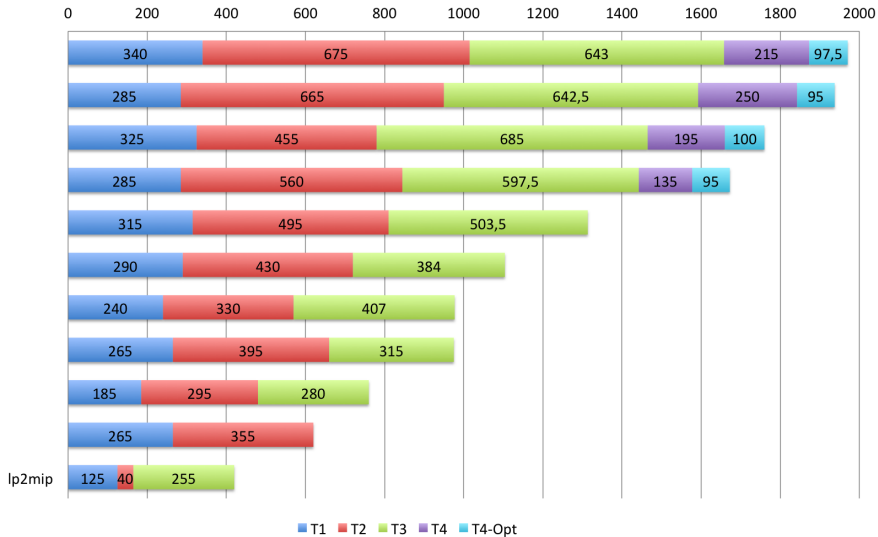
T4 Results



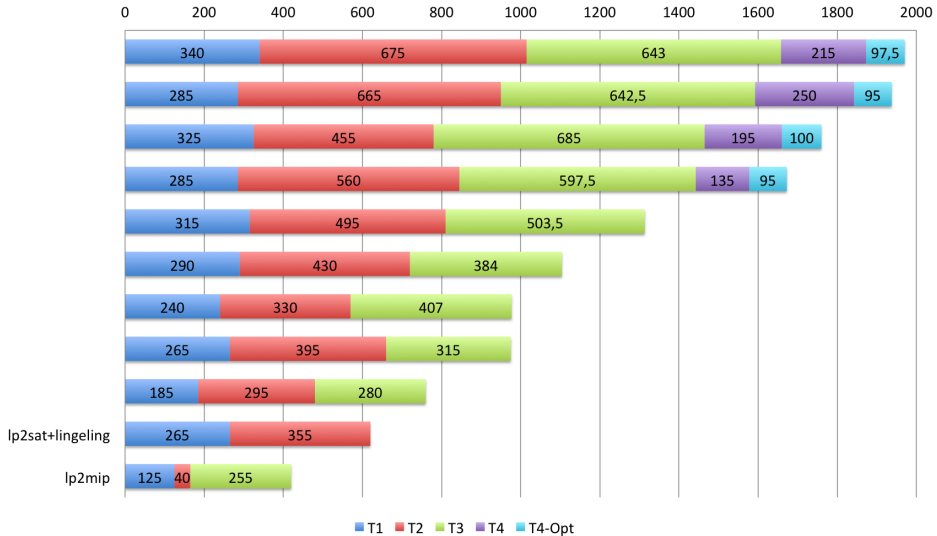
Results: ASP Competition 2014 - Overall



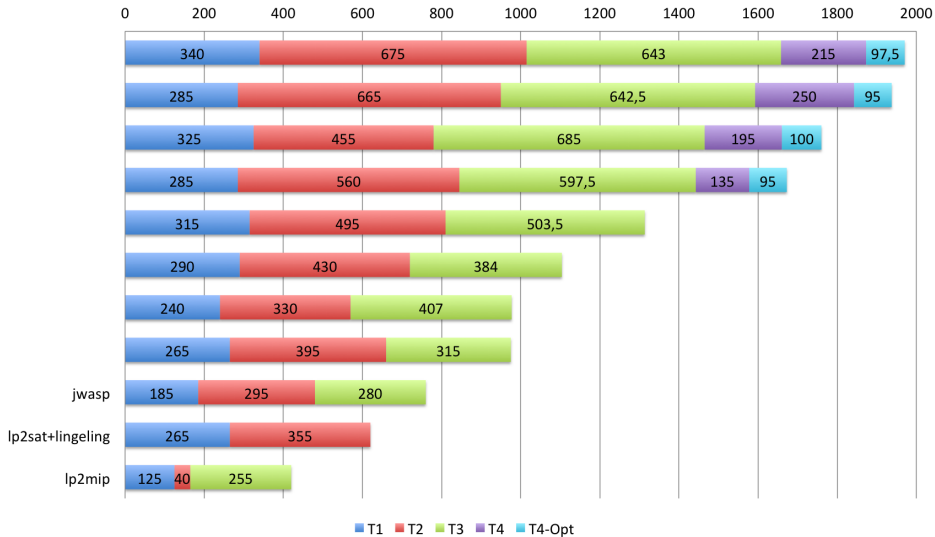
Results: ASP Competition 2014 - Overall



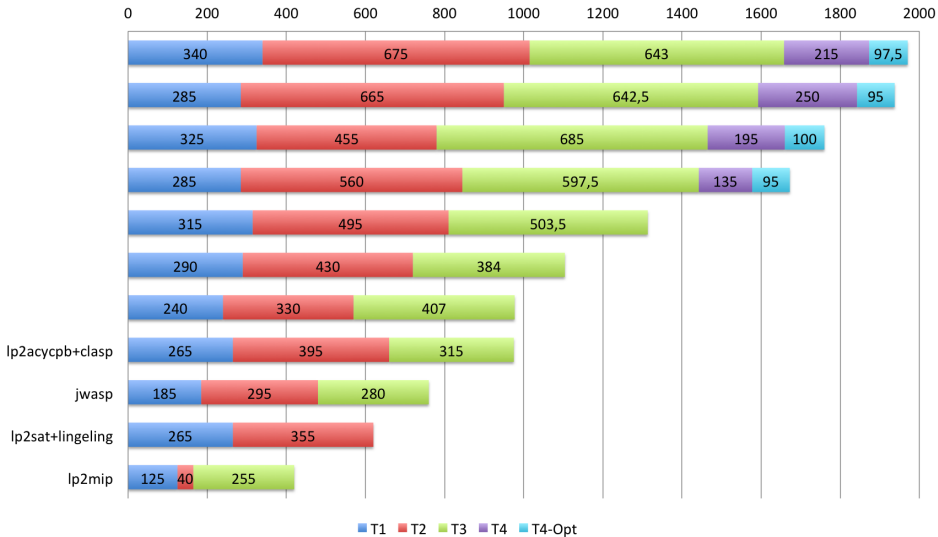
Results: ASP Competition 2014 - Overall



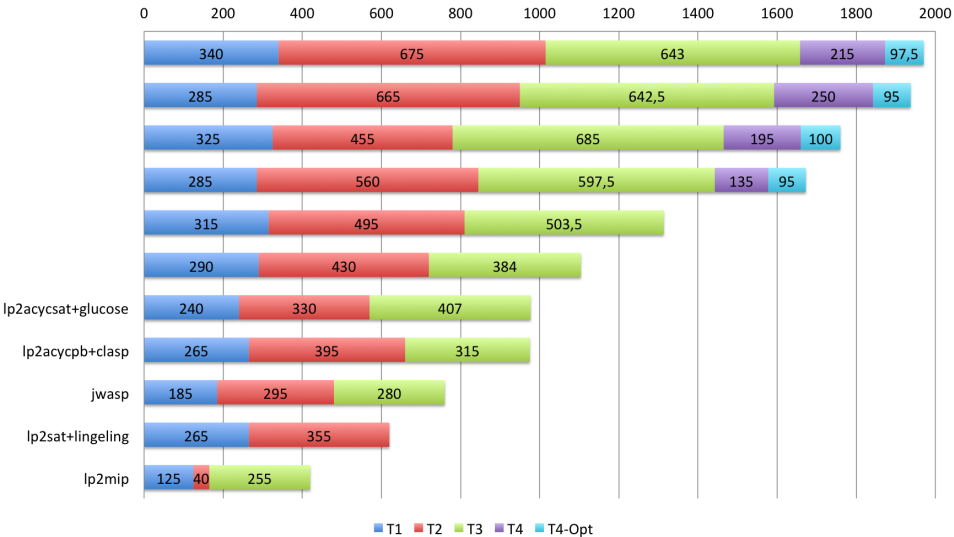
Results: ASP Competition 2014 - Overall



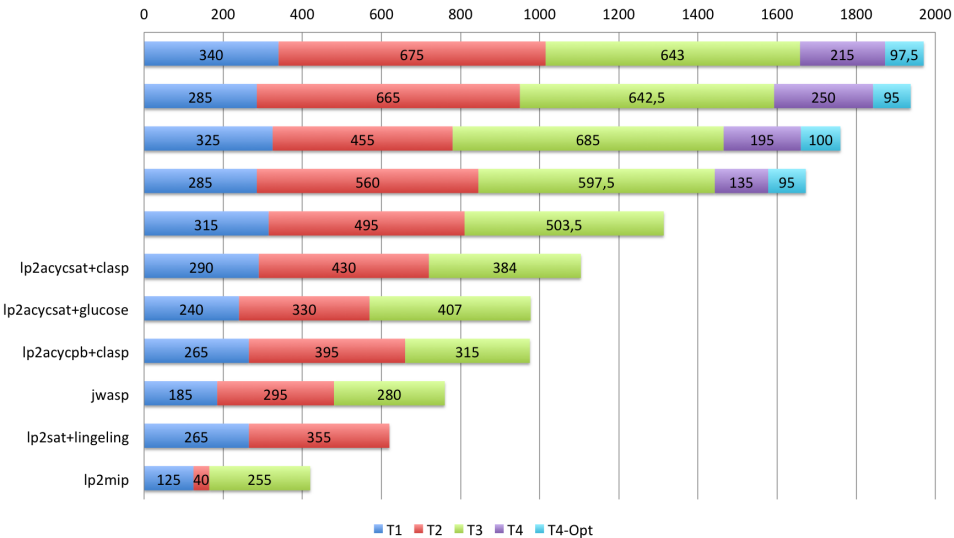
Results: ASP Competition 2014 - Overall



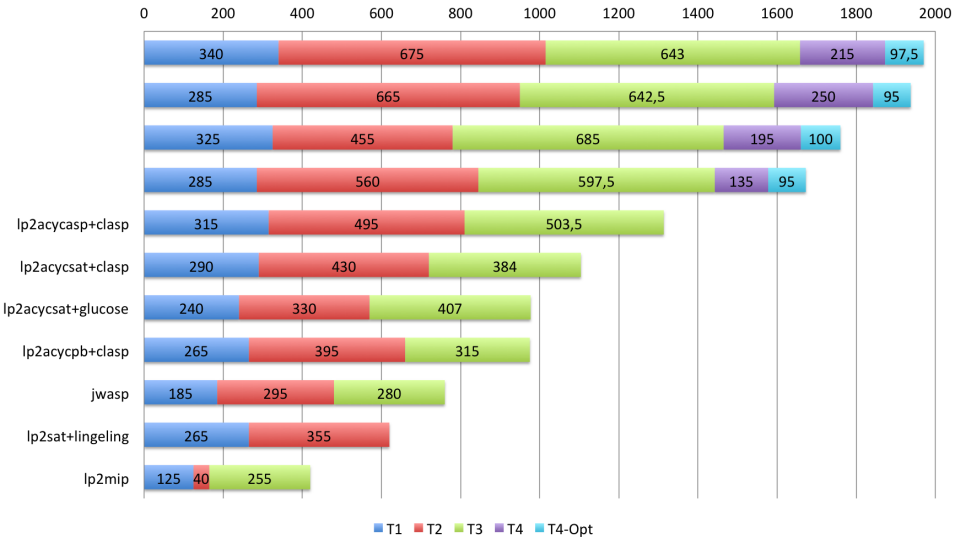
Results: ASP Competition 2014 - Overall



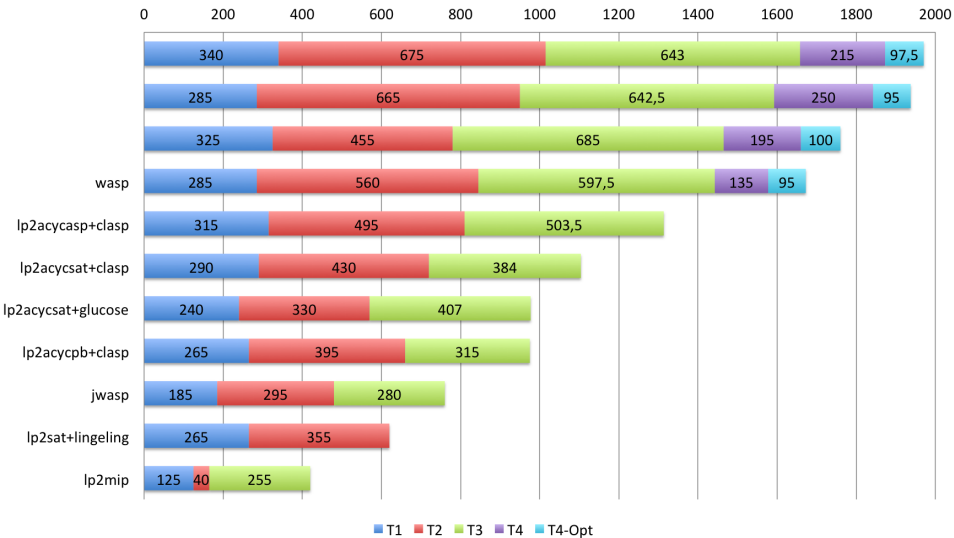
Results: ASP Competition 2014 - Overall



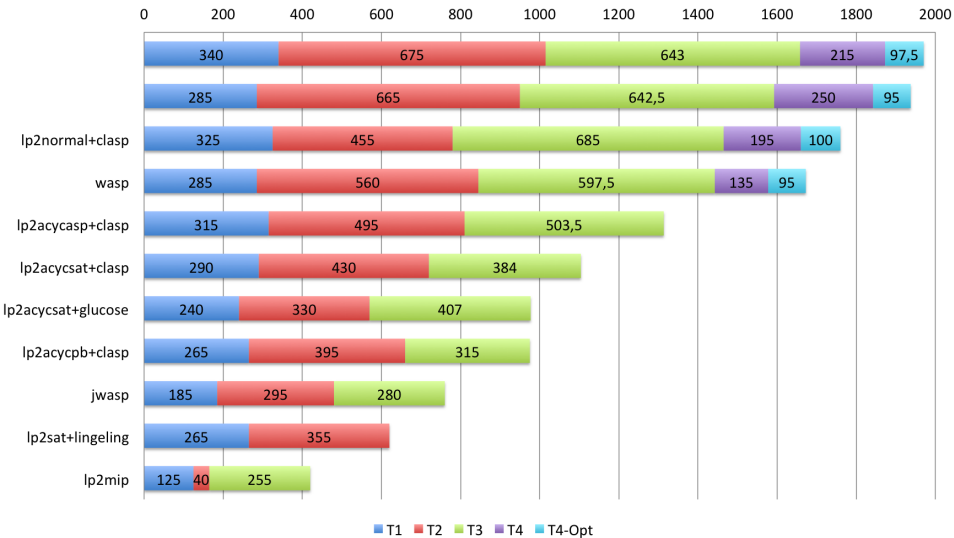
Results: ASP Competition 2014 - Overall



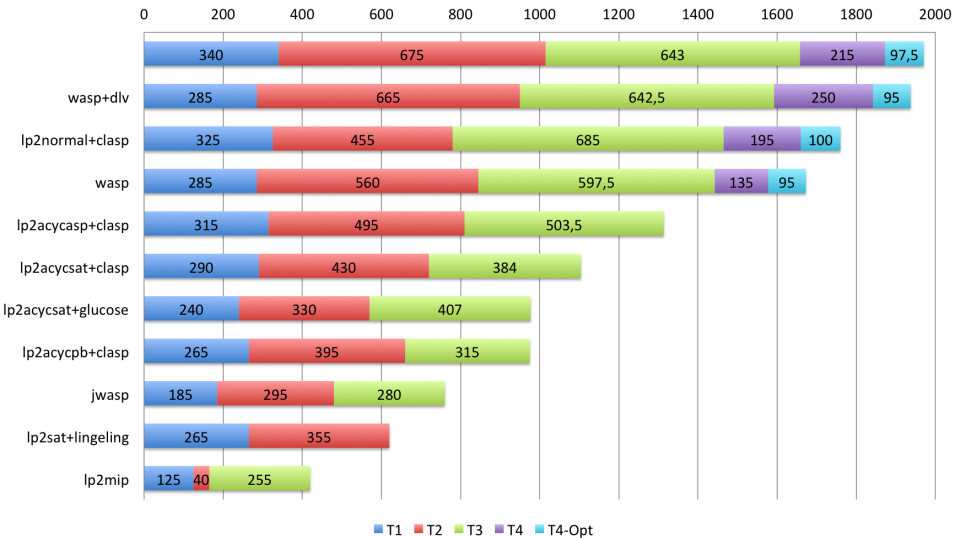
Results: ASP Competition 2014 - Overall



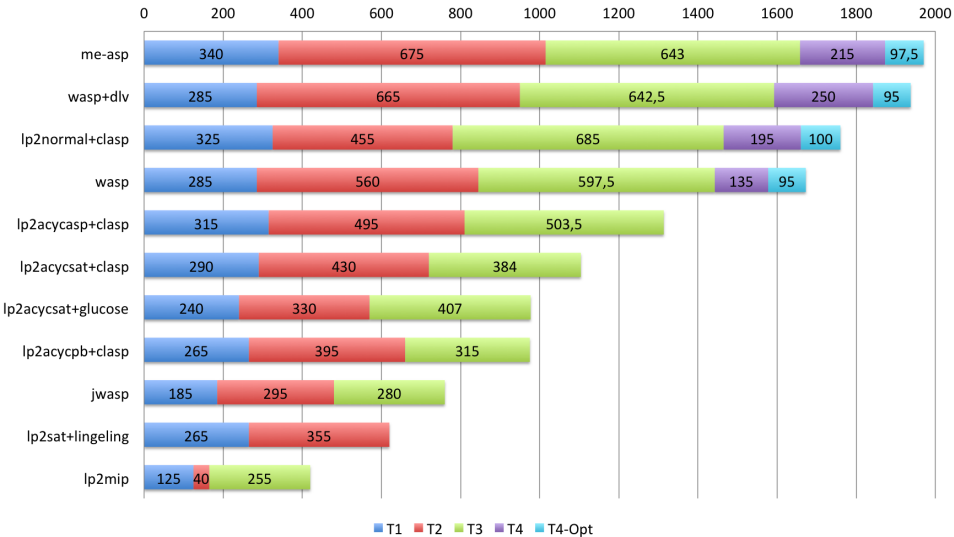
Results: ASP Competition 2014 - Overall



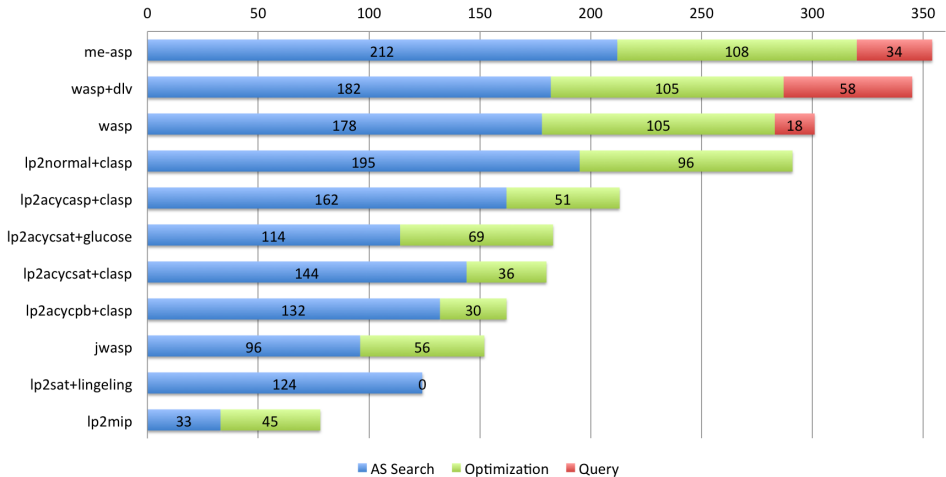
Results: ASP Competition 2014 - Overall



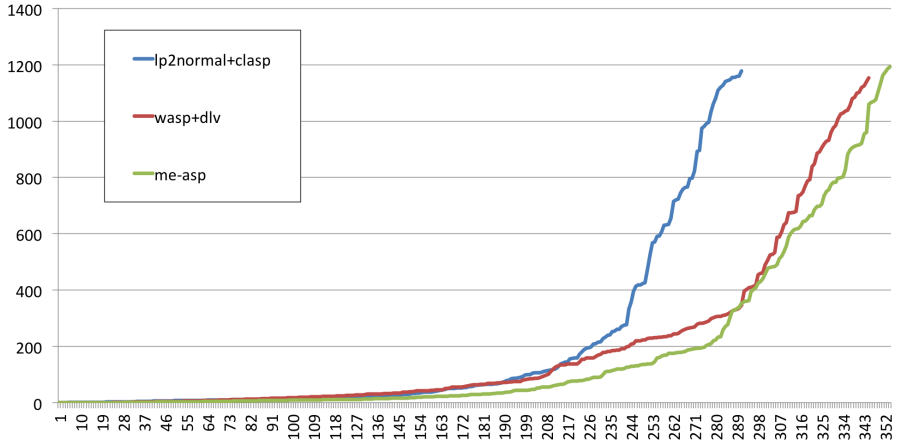
Results: ASP Competition 2014 - Overall



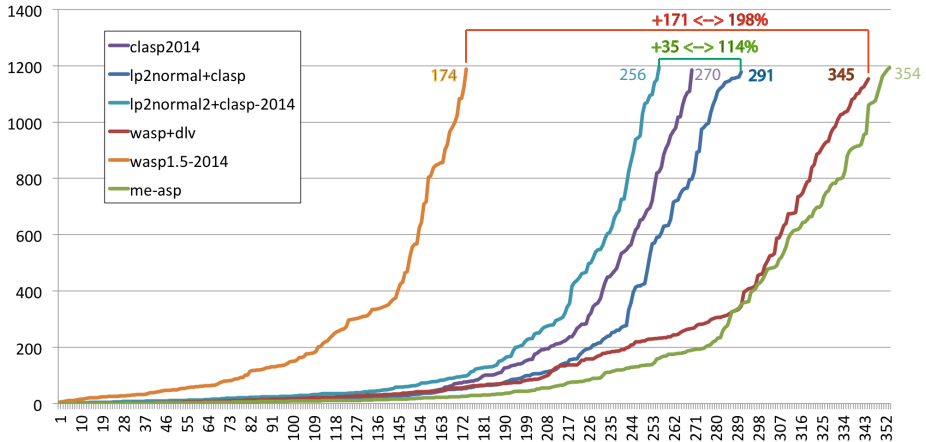
Results: Solved By Task



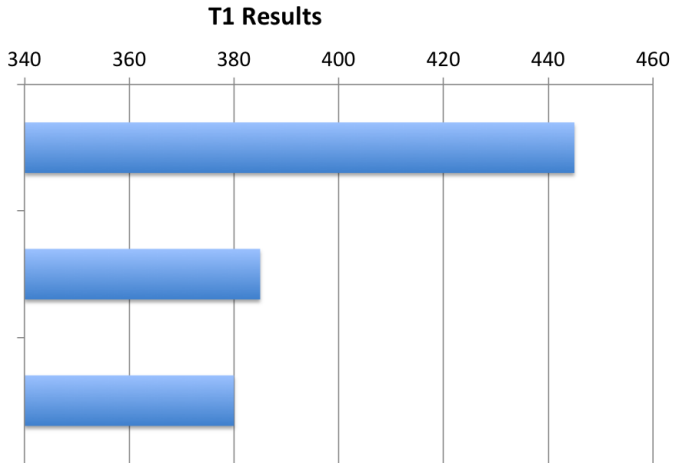
Results: Cactus Plot



Results: State of the art

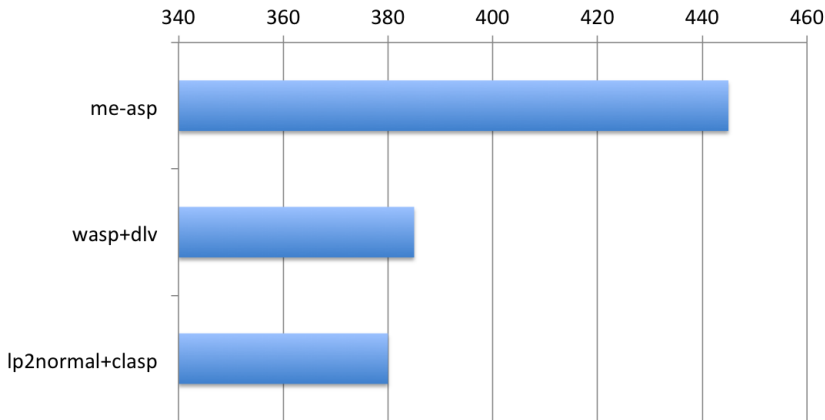


Results: Marathon Track 1 - Basic

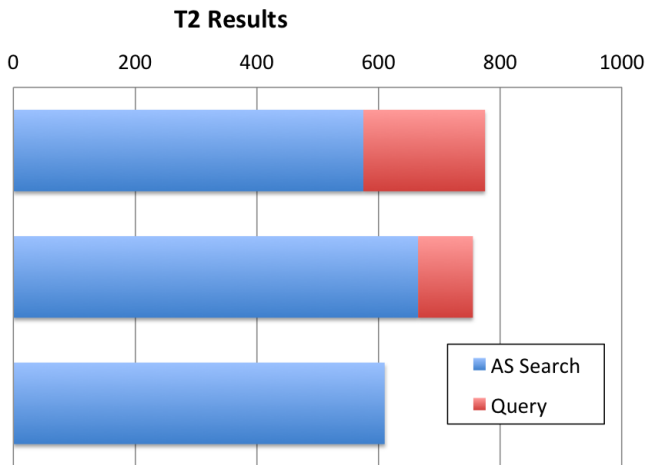


Results: Marathon Track 1 - Basic

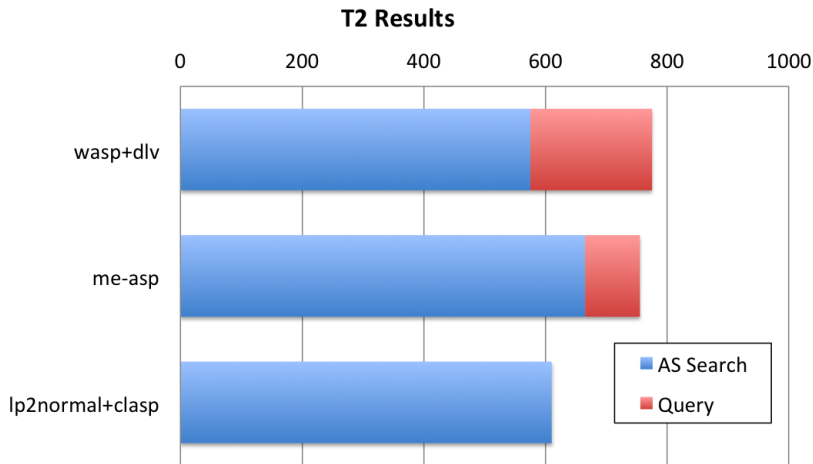
T1 Results



Results: Marathon Track 2 - Advanced

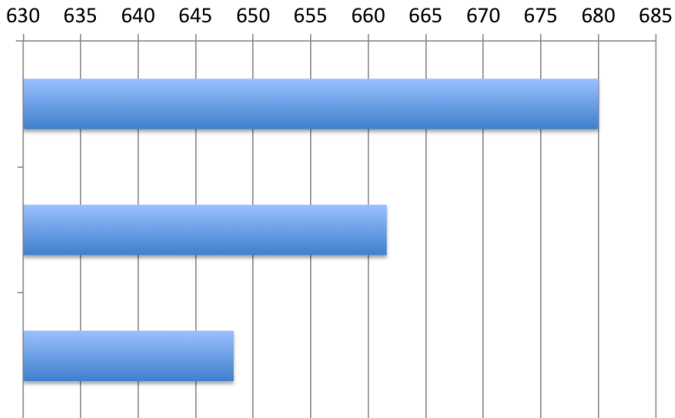


Results: Marathon Track 2 - Advanced



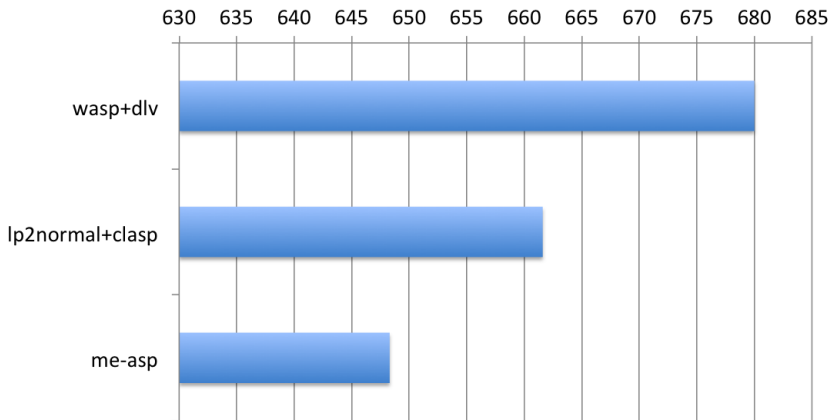
Results: Marathon Track 3 - Optimization

T3 Results

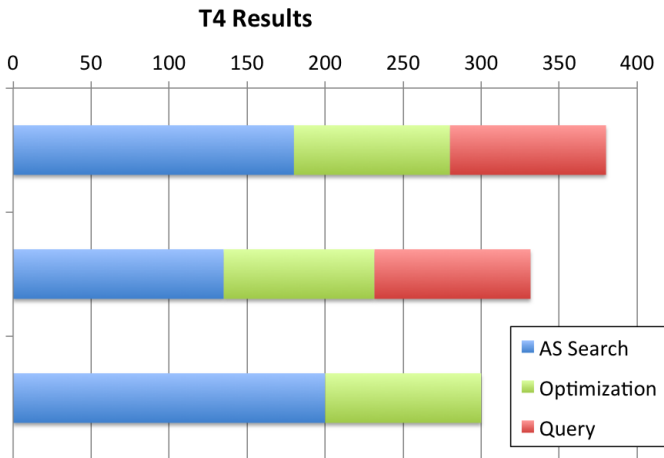


Results: Marathon Track 3 - Optimization

T3 Results

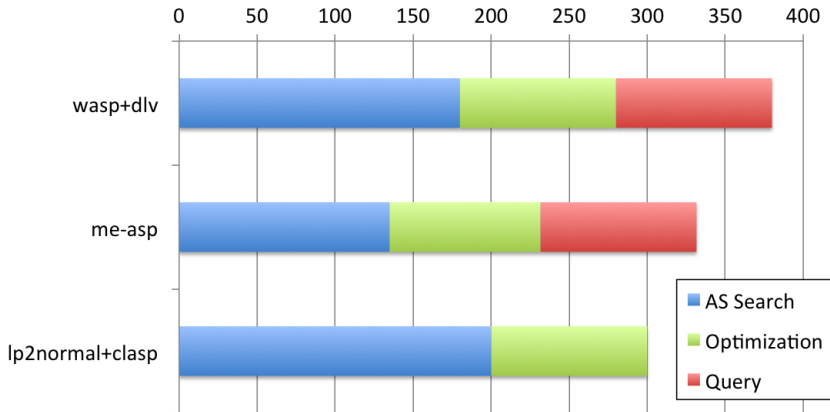


Results: Marathon Track 4 - Unrestricted

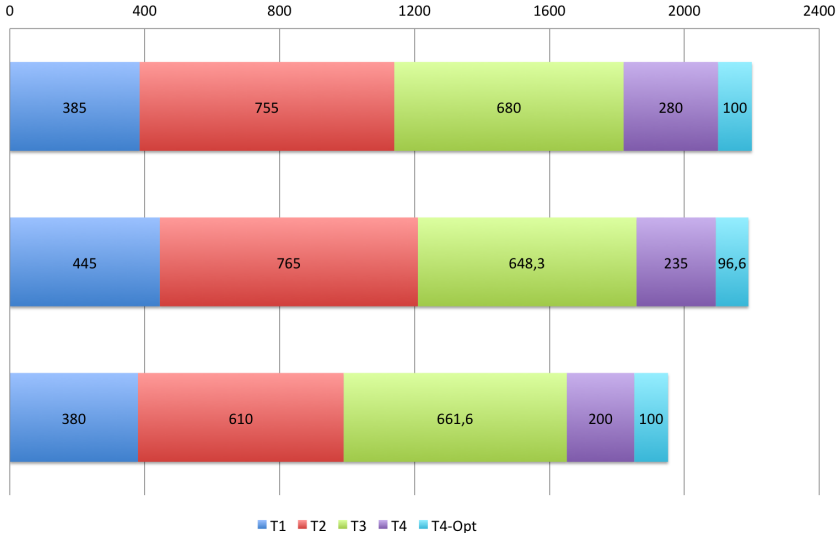


Results: Marathon Track 4 - Unrestricted

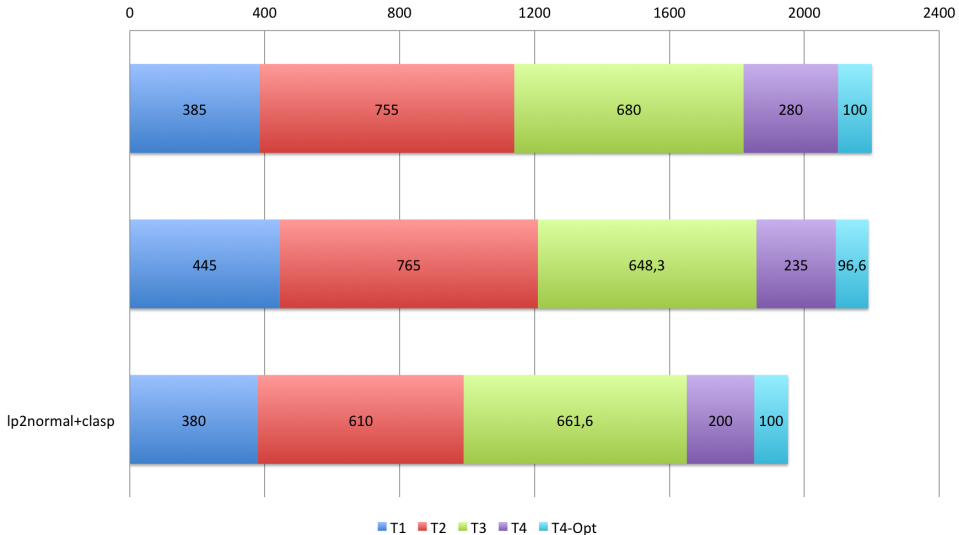
T4 Results



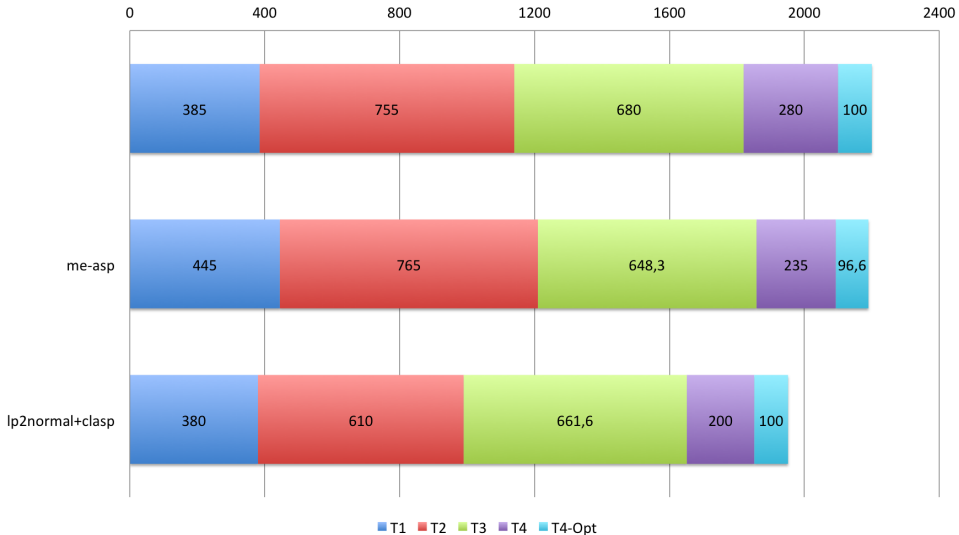
Results: Marathon ASP Competition 2015 - Overall



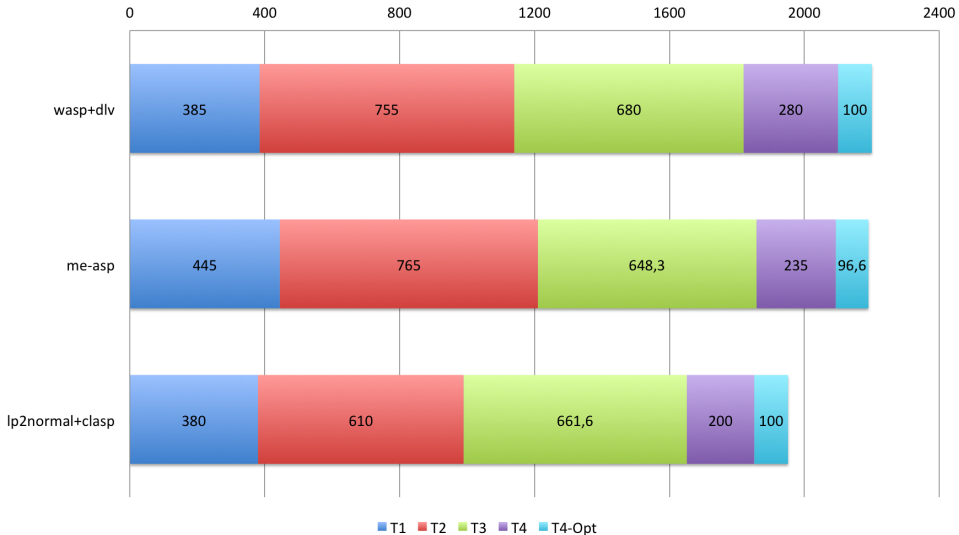
Results: Marathon ASP Competition 2015 - Overall



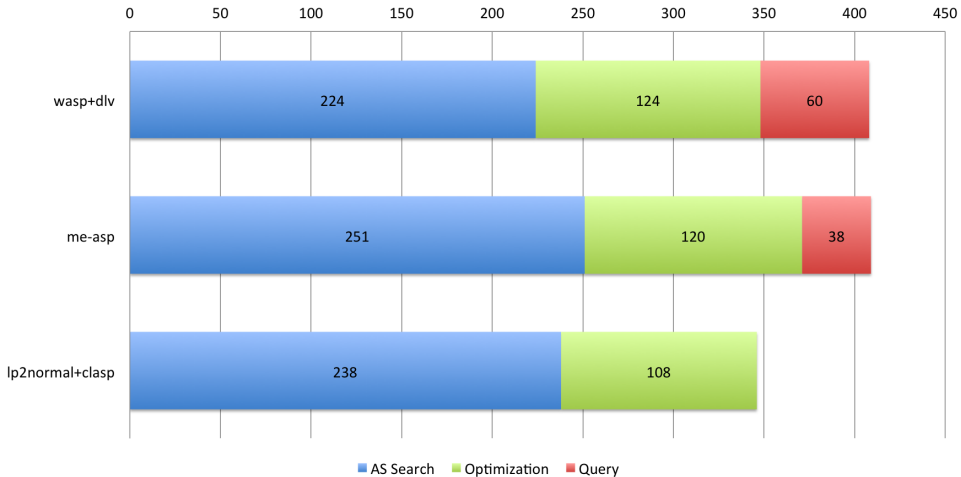
Results: Marathon ASP Competition 2015 - Overall



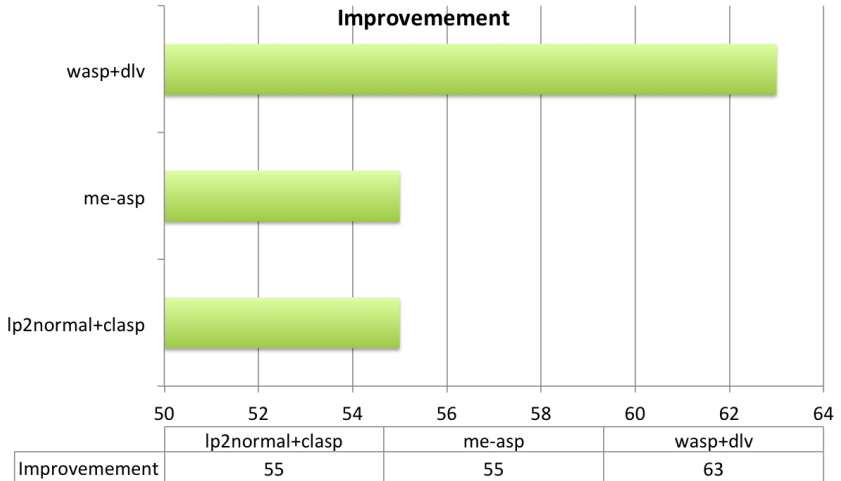
Results: Marathon ASP Competition 2015 - Overall



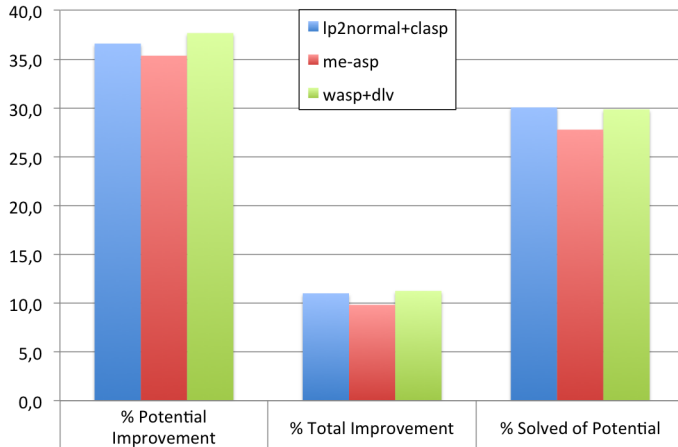
Results: Marathon Solved By Task



Results: Marathon Improvements

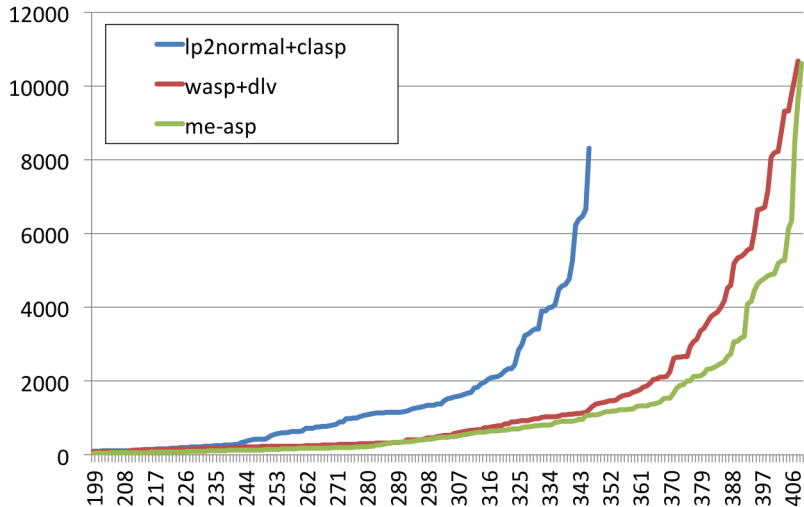


Results: Marathon Stats

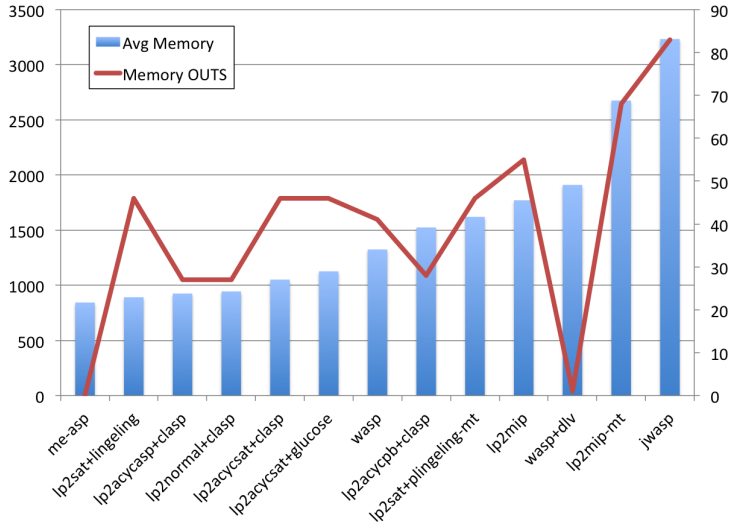


	% Potential Improvement	% Total Improvement	% Solved of Potential
lp2normal+clasp	36,6	11,0	30,1
me-asp	35,4	9,8	27,8
wasp+dlv	37,7	11,3	29,9

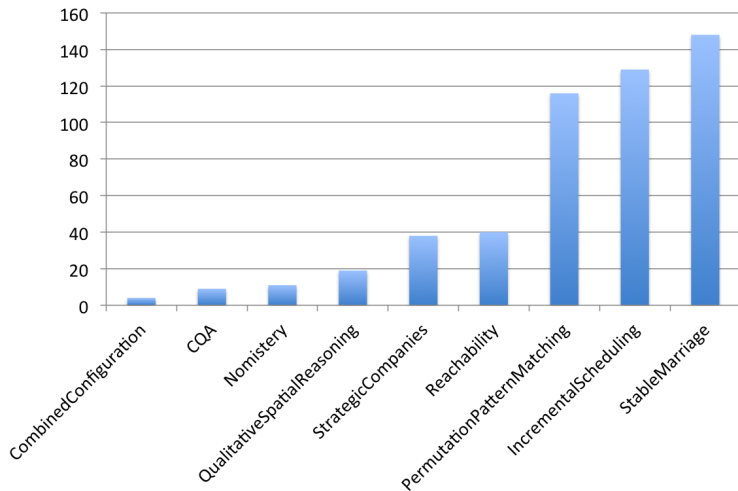
Results: Marathon Cactus Plot



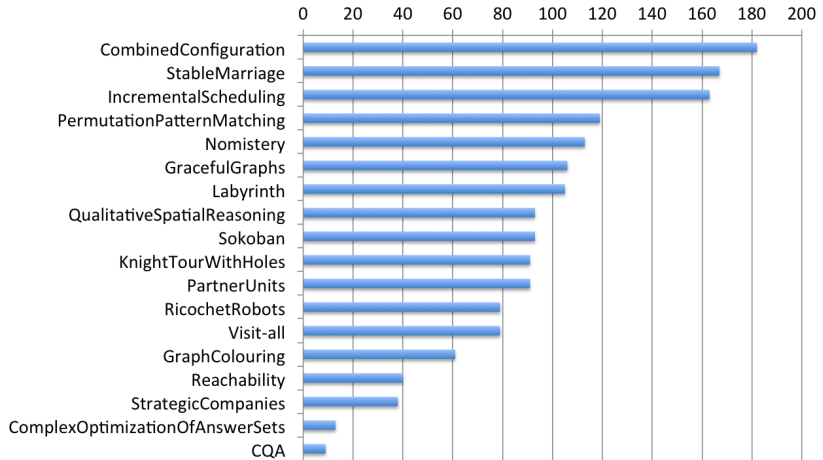
Results: Memory usage



Results: Memory out by domain



Results: Timeouts by Domain



(Implemented) Suggestions from 2014 Event

Simplify Output

- Unify output for tasks
- Reduce number of exit codes

Instance Selection

- Process for discarding very easy/hard
- More ASP-oriented real-world applications
- Enforce classification by language features
- Non-ground and ground tracks?
- Cautious/Brave Reasoning?

Modeling Competition

- Interactive event? Challenges? ...

(Implemented) Suggestions from 2014 Event

Simplify Output

- Unify output for tasks ← DONE!
- Reduce number of exit codes ← DONE!

Instance Selection

- Process for discarding very easy/hard
- More ASP-oriented real-world applications
- Enforce classification by language features
- Non-ground and ground tracks?
- Cautious/Brave Reasoning?

Modeling Competition

- Interactive event? Challenges? ...

(Implemented) Suggestions from 2014 Event

Simplify Output

- Unify output for tasks ← DONE!
- Reduce number of exit codes ← DONE!

Instance Selection

- Process for discarding very easy/hard ← DONE!
- More ASP-oriented real-world applications ← DONE!
- Enforce classification by language features ← DONE!
- Non-ground and ground tracks?
- Cautious/Brave Reasoning?

Modeling Competition

- Interactive event? Challenges? ...

(Implemented) Suggestions from 2014 Event

Simplify Output

- Unify output for tasks ← DONE!
- Reduce number of exit codes ← DONE!

Instance Selection

- Process for discarding very easy/hard ← DONE!
- More ASP-oriented real-world applications ← DONE!
- Enforce classification by language features ← DONE!
- Non-ground and ground tracks? Need for more grounders!
- Cautious/Brave Reasoning? Brave reasoning is missing

Modeling Competition

- Interactive event? Challenges? ... ideas?

Suggestions for future ASP events (1)

Simplify Output

- Avoid using exit codes with custom semantics
 - Easy choice for SAT, not for ASP solver scripts!
- Embrace POSIX-compatible convention
 - Zero for success and non-zero for error

Scoring

- Less dependent on number of participants
- More emphasis on solved (optimal) solutions
 - 5 points is too much for non-optimal witnesses
- Two rankings?

Suggestions for future ASP events (2)

Benchmark Suite

- Maintain classification by hardness
- Don't stop adding ASP-oriented real-world applications
- Maintain classification by language features
- Maintain some more easy domains

Reasoning Tasks

- Brave Reasoning
- Propositional program evaluation
- Tracks for extended language features

Suggestions for future ASP events (3)

Modeling Competition

- Keep it as it is... pure fun!
- Open to remote participation as LP/CP Contest at ICLP
- More advertisement

Extend the ASP Development community

- ASPLib web site
- Lower the entrance barrier
 - Emphasize winners of tracks...
- ... ideas?

Thanks

Thank you for your attention!

Thanks

Thank you for your attention!

Award ceremony during the social dinner!